# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-104782
(43)Date of publication of application : 21.04.1995

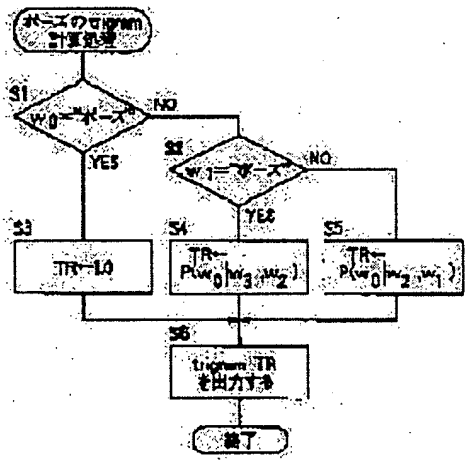
(51)Int.Cl. G10L 3/00

(21)Application number : 05-247836
(22)Date of filing : 04.10.1993

(71)Applicant : ATR ONSEI HONYAKU TSUSHIN KENKYUSHO:KK
(72)Inventor : MURAKAMI JINICHI

## (54) VOICE RECOGNITION DEVICE

(57)Abstract:
PURPOSE: To provide a voice recognition device which can obtain a high sentence recognition rate as compared with a conventional example even when an inputted spoken voice includes a pause or redundant word.
CONSTITUTION: The voice recognition device which performs voice recognition by calculating a probability value of a word to be recognized in the spoken voice sentence consisting of a character string inputted by referring to a specific statistical language model on the basis of one or plural words connected in front of the word sets the probability value of the language model of the pause in the spoken voice sentence or the word connected to the redundant word to 1 or a value close to 1, and calculates the probability of the language model of the work by skipping the pause or redundant word connected to a work other than the pause or redundant work across the pause or redundant work, thereby performing the voice recognizing process.




LEGAL STATUS

[Date of request for examination] 08.05.2000
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平7－104782

(43)公開日　平成７年(1995)４月21日

| (51)Int.Cl.[6] | 識別記号 | 庁内整理番号 | Ｆ Ｉ | 技術表示箇所 |
|---|---|---|---|---|
| Ｇ１０Ｌ　3/00 | ５３１　Ｐ | 9379－5H | | |

審査請求　未請求　請求項の数５　ＯＬ　（全７頁）

(21)出願番号　特願平5－247836

(22)出願日　平成５年(1993)10月４日

(71)出願人　593118597
株式会社エイ・ティ・アール音声翻訳通信研究所
京都府相楽郡精華町大字乾谷小字三平谷５番地

(72)発明者　村上　仁一
京都府相楽郡精華町大字乾谷小字三平谷５番地　株式会社エイ・ティ・アール音声翻訳通信研究所内

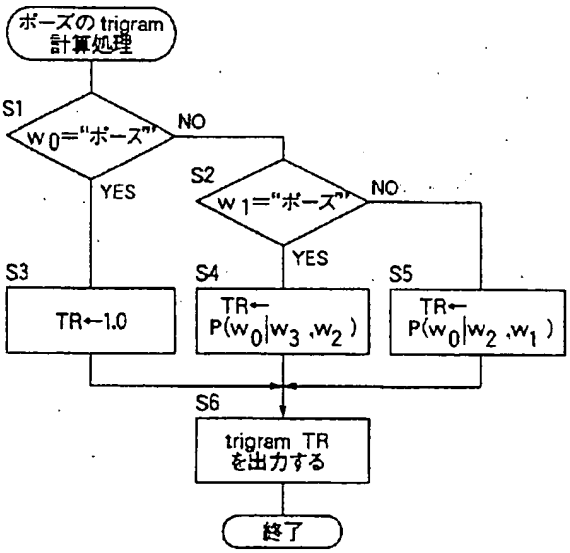
(74)代理人　弁理士　青山　葆　（外２名）

(54)【発明の名称】　音声認識装置

(57)【要約】

【目的】　入力された発声音声にポーズや冗長語があった場合であっても、従来例に比較して高い文認識率を得ることができる音声認識装置を提供する。

【構成】　所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、上記発声音声文内のポーズ又は冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズ又は冗長語を介してポーズ及び冗長語以外の単語に接続されるとき上記ポーズ又は冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する。



1

【特許請求の範囲】

【請求項１】　所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、

上記発声音声文内のポーズ又は冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズ又は冗長語を介してポーズ及び冗長語以外の単語に接続されるとき上記ポーズ又は冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備えたことを特徴とする音声認識装置。

【請求項２】　所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、

上記発声音声文内のポーズに接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズを介してポーズ以外の単語に接続されるとき上記ポーズをスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備えたことを特徴とする音声認識装置。

【請求項３】　所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、

上記発声音声文内の冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、冗長語を介して冗長語以外の単語に接続されるとき上記冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備えたことを特徴とする音声認識装置。

【請求項４】　上記所定値は、０．８以上１．０以下であることを特徴とする請求項１、２又は３記載の音声認識装置。

【請求項５】　上記統計的言語モデルは、単語のｔｒｉｇｒａｍであることを特徴とする１、２、３又は４記載の音声認識装置。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、発声音声におけるポーズ及び冗長語について処理を行う音声認識装置に関する。

【０００２】

【従来の技術】近年、連続音声認識の研究が盛んに行われ、いくつかの研究機関で文音声認識システムが構築されている。これらのシステムの多くは丁寧に発声された

2

音声を入力対象にしている。しかしながら、人間同士のコミュニケーションでは、「あのー」、「えーと」などに代表される冗長語や、一時的に発声音声が無い状態（以下、ポーズという。）である言い淀みや言い誤り及び言い直しなどが頻繁に出現する。

【０００３】

【発明が解決しようとする課題】一方、連続音声認識に利用できるアルゴリズムの１つにＯｎｅ　ｐａｓｓ　ＤＰ（Ｖｉｔｅｒｂｉ　ｓｅａｒｃｈ）アルゴリズムがある。このアルゴリズムを用いた音声認識装置においては、入力された発声音声にポーズや冗長語があった場合、文認識率が低下するという問題点があった。

【０００４】本発明の目的は以上の問題点を解決し、入力された発声音声にポーズや冗長語があった場合であっても、従来例に比較して高い文認識率を得ることができる音声認識装置を提供することにある。

【０００５】

【課題を解決するための手段】本発明に係る請求項１記載の音声認識装置は、所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、上記発声音声文内のポーズ又は冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズ又は冗長語を介してポーズ及び冗長語以外の単語に接続されるとき上記ポーズ又は冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備えたことを特徴とする。

【０００６】また、請求項２記載の音声認識装置は、所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、上記発声音声文内のポーズに接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズを介してポーズ以外の単語に接続されるとき上記ポーズをスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備えたことを特徴とする。

【０００７】さらに、請求項３記載の音声認識装置は、所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、上記発声音声文内の冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、冗長語を介して冗長語以外の単語に接続されるとき上記冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する音声認識手段を備

えたことを特徴とする。

【０００８】また、請求項４記載の音声認識装置は、請求項１、２又は３記載の音声認識装置において、上記所定値は、０．８以上１．０以下であることを特徴とする。さらに、請求項５記載の音声認識装置は、請求項１、２、３又は４記載の音声認識装置において、上記統計的言語モデルは、単語のｔｒｉｇｒａｍであることを特徴とする。

【０００９】

【作用】請求項１記載の音声認識装置においては、上記音声認識手段は、上記発声音声文内のポーズ又は冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズ又は冗長語を介してポーズ及び冗長語以外の単語に接続されるとき上記ポーズ又は冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する。

【００１０】また、請求項２記載の音声認識装置においては、上記音声認識手段は、上記発声音声文内のポーズに接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズを介してポーズ以外の単語に接続されるとき上記ポーズをスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する。

【００１１】さらに、請求項３記載の音声認識装置においては、上記音声認識手段は、上記発声音声文内の冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、冗長語を介して冗長語以外の単語に接続されるとき上記冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行する。

【００１２】

【実施例】以下、図面を参照して本発明に係る実施例の音声認識装置について説明する。図１の本実施例の音声認識装置において、音素照合部４は、入力される発声音声に関するデータに基づいて、隠れマルコフモデル（以下、HMMという。）メモリ５内の音響モデルであるHMMを参照して冗長語やポーズを認識した後、Ｏｎｅ　ｐａｓｓ　ＤＰアルゴリズムを用いたＯｎｅ　ｐａｓｓ　ＤＰ音声認識部（以下、音声認識部という。）６は、統計的言語モデルメモリ７内の統計的言語モデル（単語のｔｒｉｇｒａｍ）を参照して音声認識を実行するときに、冗長語又はポーズをスキップすることにより、冗長語及び／又はポーズを含んだ音声を認識することを特徴とする。

【００１３】まず、文音声認識アルゴリズムの改良について述べる。上記のＯｎｅ　ｐａｓｓ　ＤＰアルゴリズムの経路計算において、最尤の単語列を得るために２つの方法がある。

（１）トレースバック：各時刻・各状態において、最大累積尤度を計算したときに、選択した経路を記憶しておく。そして尤度計算の終了後、上記記憶された経路に沿ってトレースバックを行なうことによって音声認識処理を実行する（以下、第１の経路計算方法という。；例えば、Kai-Fu Lee, "Large-Vocabulary Speaker Independent Continuous Speech Recognition:The SPHINX System",15213 CMU-CS-88-148,１９８８年４月１８日参照。）。

（２）最大累積尤度と同時に計算する：各時刻、各状態において、最大累積尤度を計算したときに、同時に、選択した経路を次の状態に渡すことによって音声認識処理を実行する（以下、第２の経路計算方法という。；例えば、村上仁一，"単語のｔｒｉｇｒａｍを用いた連続音声認識の１アルゴリズム"，音響学会講演論文集，ｐｐ．１８５－１８６，２－Ｑ－７，１９９２年１０月参照。）。

【００１４】上記第１の経路計算方法は、計算量が少なくて済むため従来からしばしば利用されている。一方、上記第２の経路計算方法は、第１の経路計算方法と比較すると計算量が増加するが、各時刻・各状態においてトレースバックをしなくても経路を知ることが可能であるため、言語モデルにおけるｌｅｆｔ－ｒｉｇｈｔ型のＬＲパーザと組み合わせることが容易であり、また、多くの場合、第１の経路計算方法と比較して少量のメモリですむので、本実施例の音声認識部６は後者の第２の経路計算方法を採用する。

【００１５】以下、本実施例の音声認識方法を用いる音声認識装置を示す図１を参照して、本実施例の統計的言語モデルを用いた音声認識装置の構成及び動作について説明する。

【００１６】図１において、話者の発声音声はマイクロホン１に入力されて音声信号に変換された後、特徴抽出部２に入力される。特徴抽出部２は、入力された音声信号をＡ／Ｄ変換した後、例えばＬＰＣ分析を実行し、対数パワー、１６次ケプストラム係数、Δ対数パワー及び１６次Δケプストラム係数を含む３４次元の特徴パラメータを抽出する。抽出された特徴パラメータの時系列はバッファメモリ３を介して音素照合部４に入力される。音素照合部４に接続される隠れマルコフモデル（以下、HMMという。）メモリ５内のHMMは、複数の状態と、各状態間の遷移を示す弧から構成され、各弧には状態間の遷移確率と入力コードに対する出力確率を有している。音素照合部４は、入力されたデータに基づいて音素照合処理を実行して音素データを、音声認識部６に出力する。

【００１７】単語のｔｒｉｇｒａｍを含む所定の統計的言語モデルを予め記憶する統計的言語モデルメモリ７は音声認識部６に接続される。音声認識部６は、統計的言語モデルメモリ７内の統計的言語モデルを参照して、所定のＯｎｅ　ｐａｓｓ　ＤＰアルゴリズムを用いて、入力された音素データについて左から右方向に、後戻りな

5

しに処理してより高い生起確率の単語を音声認識結果データと決定することにより音声認識の処理を実行して、決定された音声認識結果データ（文字列データ）を出力する。

【００１８】次いで、音声認識部６におけるポーズの処理について説明する。ポーズは文節間において出現することが多いが、音声のあらゆる場所に出現する可能性がある。しかしながら、従来の言語モデルでは、これに追随できないため、ポーズの区間で誤認識が起きやすい。そこで、単語のｔｒｉｇｒａｍなる統計的言語モデルと、Ｏｎｅ　ｐａｓｓ　ＤＰアルゴリズムを用いて、全ての単語と単語の境界にポーズが入力されても文認識が可能な音声認識方法を発明した。本発明に係る実施例においては、まずポーズを１単語と考えて、ポーズに接続される単語のｔｒｉｇｒａｍの値は１．０にする。そしてポーズ以外の単語に接続されるときポーズをスキップして単語のｔｒｉｇｒａｍを計算する。例えば「東京都　港区　新橋／ｐａｕｓｅ（ポーズ）／１丁目」という文字列なる文が入力されたとき、その確率を、Ｐ（新橋｜東京都　港区）×１．０×Ｐ（１丁目｜港区　新橋）と計算する。ここで、Ｐ（Ａ｜Ｂ）は、単語Ｂの後に単語Ａが来る確率であり、以下、同様である。このようにすると、近似解ではあるが、ポーズを除いて単語ｔｒｉｇｒａｍを用いたときの最尤の解が得られる。

【００１９】以上の実施例において、ポーズに接続される単語のｔｒｉｇｒａｍの値を１．０としているが、本発明はこれに限らず、ポーズに接続される単語のｔｒｉｇｒａｍの値を好ましくは０．８以上１．０以下の範囲のある値に設定する。

【００２０】図２は、図１の音声認識装置において実行されるポーズのｔｒｉｇｒａｍ計算処理を示すフローチャートである。なお、冗長語についても図２のフローと同様に処理される。当該計算処理は、単語列$w_3$，$w_2$，$w_1$，$w_0$が入力されたときに単語$w_0$のｔｒｉｇｒａｍを計算する方法であって、図２に示すように、まず、ステップＳ１において単語$w_0$がポーズであるか否かが判断され、ステップＳ２において、タンゴ$w_1$がポーズであるか否かが判断される。ステップＳ１においてＹＥＳであれば、ステップＳ３において単語$w_0$のｔｒｉｇｒａｍの値ＴＲを１．０に設定した後、ステップＳ６に進む。ステップＳ１においてＮＯの場合は、ステップＳ２に進み、ステップＳ２においてＹＥＳであれば、ステップＳ４において確率値Ｐ（$w_0$｜$w_3$，$w_2$）を単語$w_0$のｔｒｉｇｒａｍの値ＴＲとして設定した後、ステップＳ６に進む。一方、ステップＳ２においてＮＯであれば、ステップＳ５において確率値Ｐ（$w_0$｜$w_2$，$w_1$）を単語$w_0$のｔｒｉｇｒａｍの値ＴＲとして設定した後、ステップＳ６に進む。ステップＳ６においては、設定された単語$w_0$のｔｒｉｇｒａｍの値ＴＲを計算値として出力して当該計算処理を終了する。

6

【００２１】さらに、音声認識部６における冗長語の処理について説明する。自由発話では「あのー」、「えーと」などの冗長語が多く出現する。本発明者が、詳細後述するテストデータから収集した自由発話における出現回数が２以上の冗長語を表１乃至表３に示す。

【００２２】

【表１】

| 冗長語 | 出現回数 |
|---|---|
| 「あ」 | ６０４ |
| 「あー」 | ２６８ |
| 「あーっと」 | ２ |
| 「あーん」 | ５ |
| 「ああ」 | ７ |
| 「あっ」 | １５１ |
| 「あの」 | １８０９ |
| 「あのー」 | ２０２５ |
| 「あのう」 | ７７ |
| 「あのうー」 | ３ |
| 「い」 | ２６ |
| 「いー」 | ５８ |
| 「いやー」 | ２ |
| 「う」 | ２３ |
| 「うー」 | ７１ |
| 「うーん」 | ２６ |
| 「うーんと」 | ２ |
| 「うん」 | ７ |
| 「え」 | １０４０ |
| 「えー」 | ３１０５ |
| 「えーっと」 | ２５６ |
| 「えーっとー」 | ２ |

【００２３】

【表２】

| 冗長語 | 出現回数 |
|---|---|
| 「えーっとですね」 | ８ |
| 「えーと」 | ４６６ |
| 「えーとー」 | ４ |
| 「えーとですね」 | ３ |
| 「えーまあ」 | ３ |
| 「えーん」 | ２ |
| 「ええ」 | １３ |
| 「えっ」 | ２２ |
| 「えっーと」 | ４ |
| 「えっと」 | ６２ |
| 「えっとー」 | １１ |
| 「えと」 | ４７ |

| | |
|---|---|
| 「えとー」 | １３ |
| 「お」 | ５９ |
| 「おー」 | １９６ |
| 「おっ」 | ２ |
| 「こう」 | ９ |
| 「この」 | ９ |
| 「このー」 | ４ |
| 「じゃ」 | ４ |
| 「す」 | ８ |
| 「すー」 | ２ |

【００２４】
【表３】

| 冗長語 | 出現回数 |
|---|---|
| 「すっ」 | ２ |
| 「そ」 | ２ |
| 「その」 | １１５ |
| 「そのー」 | ４８ |
| 「ちょっと」 | ８ |
| 「つ」 | ２ |
| 「で」 | ６１ |
| 「でー」 | １３ |
| 「と」 | ７７ |
| 「とー」 | １１ |
| 「は」 | ４ |
| 「はあー」 | ２ |
| 「ふーん」 | ２ |
| 「ま」 | ２６３ |
| 「まー」 | ８ |
| 「まぁ」 | １８６ |
| 「まあ」 | １７６ |
| 「まっ」 | ５ |
| 「も」 | ２ |
| 「ん」 | ２７ |
| 「んー」 | １９ |
| 「んと」 | ２ |

【００２５】この冗長語は文の全ての場所に出現する可能性があるという点でポーズと似た性質がある。従って、冗長語の処理にはポーズの処理と同様な手法が使用できる。すなわち、音響モデルでは冗長語を認識しながら、言語モデルでは冗長語をスキップする。例えば「東京都　港区　新橋　えーと　１丁目」という文字列なる文が入力されたとき、その確率値を、Ｐ（新橋｜　東京都　港区）×１．０×Ｐ（１丁目｜港区　新橋）と計算する。従って、冗長語の処理においても、図２に示したフローチャートにおける「ポーズ」を「冗長語」と置き換えて同様に処理することができる。



【００２６】以上の実施例において、冗長語に接続される単語のｔｒｉｇｒａｍの値を１．０としているが、本発明はこれに限らず、冗長語に接続される単語のｔｒｉｇｒａｍの値を好ましくは０．８以上１．０以下の範囲のある値に設定する。

【００２７】本発明者は、以上説明した本実施例の音声認識装置を用いて文認識率による評価を行うためにシミュレーションを行った。言語情報には、単語のｂｉｇｒａｍ又はｔｒｉｇｒａｍを用いて特定話者及び不特定話者の音声認識シミュレーションを行った。テストデータはナレータが発声した国際会議の問い合わせの文（いわゆる、モデル会話。）を使用した。なお、テストデータの先頭及び末尾には約２０ミリ秒のポーズがある。また、ｔｒｉｇｒａｍの連鎖確率値は、本出願人の対話データベースの中から国際会議の予約に関するデータである約１２，０００文章（約１７０，０００単語）に、テストデータのテキストを加えて計算した。

【００２８】上述のポーズの処理を行わないシミュレーションにおいては、ｔｒｉｇｒａｍの場合、特定話者では７８．６％の文認識率が得られる一方、不特定話者では５９．５％の文認識率が得られた。これに対して、図２に示したポーズの処理を実行した文認識率のシミュレーションにおいては、特定話者認識において８６．３％のより高い文認識率が得られ、音声認識性能が向上した。一方、不特定話者認識においては６０．３％の文認識率が得られ、音声認識性能の向上は顕著ではないが、若干向上した。

【００２９】さらに、自由発話の文認識シミュレーションについて述べる。ここでは、特に、冗長語の処理の有効性を自由発話の音声で調べる。自由発話の定義は人によって異なるが、本実施例を用いたシミュレーションでは以下に示すような方法で収録した音声データを使用する。

（１）朗読音声データ：テキストを読みあげた音声データであって、冗長語や、言い淀み、言い直しは無い。

（２）疑似自由発話データ：冗長語をつけてテキストを読み上げた音声データである。冗長語を除いて、上記（１）朗読発声データと発話内容は同一であって、言い淀み、言い直しは無い。

（３）自由発話データ：話者はテキストを覚えて、その意図を理解し、自由に発話した音声データである。発話内容は上記（１）朗読発声データと異なる。ここで、言い淀み、言い直しは無い。

【００３０】話者は、ナレータではなく、一般の人である。冗長語として、「あのー」「えーと」「まあ」などの１０９種類を定義する。音声認識シミュレーションは、不特定話者認識の単語のｔｒｉｇｒａｍの場合のみ行った。また、上述のポーズの処理も行った。このシミュレーション結果を表４に示す。表４から明らかなように、冗長語を付けた疑似自由発話の音声では、６４．４

%述べる認識率が得られる一方、自由発話音声では３４．４%の認識率が得られた。

＊【００３１】

＊　【表４】

自由発話の文認識シミュレーション結果－－－文認識率（%）

| 発話様式 | ＋ポーズ処理 | ＋ポーズ処理＋冗長語処理 |
|---|---|---|
| 朗読発声 | ８２．６% | ７４．８% |
| 疑似自由発話 | ２６．７% | ６４．４% |
| 自由発話 | １４．１% | ３４．４% |

【００３２】従って、上述の冗長語の処理は、疑似自由発話でも自由発話でも、ポーズの処理のみの方法より有効であること、並びに、朗読発声の文認識においても従来の方法と比較して認識率があまり低下しないこと（８２．６%→７４．８%）を考慮すると、自由発話の認識において有効であることがわかる。

【００３３】以上の実施例において、音声認識部６における冗長語の処理及びポーズの処理について説明しているが、音声認識部６は、冗長語の処理とポーズの処理とのうち少なくとも一方を含むように構成してもよい。

【００３４】

【発明の効果】以上詳述したように本発明によれば、所定の統計的言語モデルを参照して入力された文字列からなる発声音声文を、音声認識すべき単語の確率値をその単語の前に接続される１個又は複数個の単語に基づいて計算することによって音声認識する音声認識装置において、上記発声音声文内のポーズ又は冗長語に接続される単語の言語モデルの確率値を１又は１に近い値である所定値とする一方、ポーズ又は冗長語を介してポーズ及び冗長語以外の単語に接続されるとき上記ポーズ又は冗長語をスキップして単語の言語モデルの確率値を計算して音声認識処理を実行するように構成したので、ポーズ又は冗長語を考慮して音声認識することができ、これによって、認識率が大幅に向上し、音声認識性能を向上させることができる。

【図面の簡単な説明】

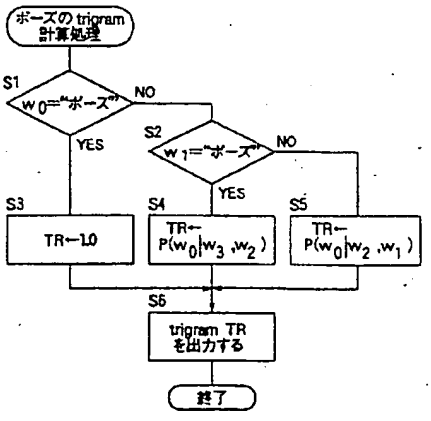
【図１】　本発明に係る一実施例である音声認識装置のブロック図である。

【図２】　図１の音声認識装置において実行されるポーズのｔｒｉｇｒａｍ計算処理を示すフローチャートである。

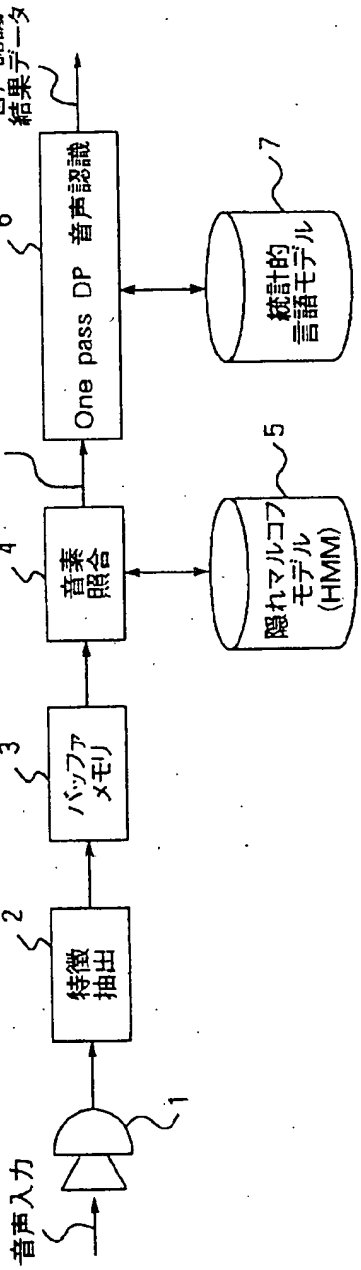
【符号の説明】

１…マイクロホン、
２…特徴抽出部、
３…バッファメモリ、
４…音素照合部、
５…Ｏｎｅ　ｐａｓｓ　ＤＰ音声認識部、
６…隠れマルコフモデル（ＨＭＭ）メモリ、
７…統計的言語モデルメモリ。

【図２】



【図1】


音声入力
1
特徴抽出
2
バッファメモリ
3
音素照合
4
隠れマルコフモデル(HMM)
5
音素データ
One pass DP 音声認識
6
統計的言語モデル
7
音声認識結果データ